

Xinya Du

CONTACT

Assistant Professor
Department of Computer Science
The University of Texas at Dallas
800 West Campbell Road, Richardson, TX 75080

Email: xinya.du@utdallas.edu
Website: <https://xinyadu.github.io>
Google Scholar

RESEARCH
INTERESTS

Natural Language Processing, Deep Learning, Large Language Models

EDUCATION

Cornell University
Ph.D. in Computer Science (M.S. degree granted in Aug 2019)
Dissertation: Towards More Intelligent Extraction of Information from Documents
Advisor: Claire Cardie

Aug 2016 – Aug 2021

Shanghai Jiao Tong University
B.E. in Computer Science and Engineering

Sep 2012 – Aug 2016

EXPERIENCE

University of Texas at Dallas, Richardson, TX
Assistant Professor in Computer Science

Aug 2022 – Present

University of Illinois at Urbana-Champaign, Champaign, IL
Postdoctoral Research Associate, with Prof. Heng Ji

Sep 2021 – Aug 2022

Google AI, Mountain View, CA
Research Intern

May 2020 – Aug 2020

Allen Institute for Artificial Intelligence, Seattle, WA
Research Intern

Sep 2018 – Dec 2018

Microsoft Research, Redmond, WA
Research Intern

May 2018 – Aug 2018

SELECTED
AWARDS&
HONORS

NSF CAREER award

National Science Foundation, 2024

AAAI New Faculty Highlights

AAAI, 2024

Amazon Research Award

Amazon, 2023

Spotlight Rising Star in Data Science

University of Chicago, 2021

Top 100 New Stars in Artificial Intelligence

Baidu Scholar, 2020

Most Influential ACL Papers (15 each year)

Paperdigest, 2017

National Scholarship (Top 1% students nationwide)

SJTU, 2013

PUBLICATIONS

Note: † indicates that I am a co-leading author. * indicates equal contributions.

[1] **FaithScore: Fine-grained Evaluations of Hallucinations in Large Vision-Language Models**
Liqiang Jing, Ruosen Li, Yunmo Chen, **Xinya Du**
In Findings of the Association for Computational Linguistics: (EMNLP), 2024.

[2] **Document-level Causal Relation Extraction with Knowledge-guided Binary Question Answering**

Zimu Wang, Lei Xia, Wei Wang, **Xinya Du**
In Findings of the Association for Computational Linguistics: (EMNLP), 2024.

- [3] **QAEvent: Event Extraction as Question-Answer Pairs Generation**
Milind Choudhary, **Xinya Du**
In Findings of the Association for Computational Linguistics: (EACL), 2024.
- [4] **PRD: Peer Rank and Discussion Improve Large Language Model based Evaluations**
Ruosen Li, Teerth Patel, Xinya Du
In Transactions on Machine Learning Research (TMLR), 2024.
- [5] **Large Language Models for Automated Open-domain Scientific Hypotheses Discovery**
Zonglin Yang, **Xinya Du**[†], Junxian Li, Jie Zheng, Soujanya Poria, Erik Cambria
In Findings of the Association for Computational Linguistics: (ACL), 2024.
Best poster award in AI4Science workshop, 2024 (1/200).
- [6] **Language Models as Inductive Reasoners**
Zonglin Yang, Li Dong, **Xinya Du**[†], Hao Cheng, Erik Cambria, Xiaodong Liu, Jianfeng Gao, Furu Wei
In Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2024.
- [7] **Leveraging Structured Information for Explainable Multi-hop Question Answering and Reasoning**
Ruosen Li, **Xinya Du**
In Findings of the Association for Computational Linguistics: (EMNLP), 2023.
- [8] **Process of Elimination for Multiple Choice Reasoning**
Chenkai Ma, **Xinya Du**
In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023.
- [9] **Probing Representations for Document-level Event Extraction**
Barry Wang, **Xinya Du**, Claire Cardie
In Findings of the Association for Computational Linguistics: (EMNLP), 2023.
- [10] **Zero-Shot Classification by Logical Reasoning on Natural Language Explanations**
Chi Han, Hengzhi Pei, **Xinya Du**, Heng Ji
In Findings of the Association for Computational Linguistics: (ACL), 2023.
- [11] **Toward Consistent and Informative Event-Event Temporal Relation Extraction**
Xiaomeng Jin, Haoyang Wen, **Xinya Du**, Heng Ji
In MATCHING at Annual Meeting of the Association for Computational Linguistics (ACL), 2023.
- [12] **End-to-end Case-Based Reasoning for Commonsense Knowledge Base Completion**
Zonglin Yang, **Xinya Du**[†], Erik Cambria, Claire Cardie
In Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2023.
- [13] **Logical Entity Representation in Knowledge-Graphs for Differentiable Rule Learning**

Chi Han, Qizheng He, Charles Yu, **Xinya Du**, Hanghang Tong, Heng Ji
In International Conference on Learning Representations (ICLR), 2023.

- [14] **Retrieval-Augmented Generative Question Answering for Event Argument Extraction**
Xinya Du and Heng Ji
In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2022.
- [15] **Dynamic Global Memory for Document-level Argument Extraction**
Xinya Du, Sha Li, and Heng Ji
In Annual Meeting of the Association for Computational Linguistics (ACL), 2022.
- [16] **Automatic Error Analysis for Document-level Information Extraction**
Aliva Das*, **Xinya Du***, Barry Wang*, Kejian Shi, Jiayuan Gu, Thomas Porter, Claire Cardie
In Annual Meeting of the Association for Computational Linguistics (ACL), 2022.
- [17] **RESIN-11: Schema-guided Event Prediction for 11 Newsworthy Scenarios**
Xinya Du, Zixuan Zhang, Sha Li, Heng Ji and the RESIN team
In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL): System Demonstrations, 2022.
- [18] **Template Filling with Generative Transformers**
Xinya Du, Alexander M. Rush, and Claire Cardie
In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2021.
- [19] **GRIT: Generative Role-filler Transformers for Document-level Event Entity Extraction**
Xinya Du, Alexander M. Rush, and Claire Cardie
In Conference of the European Chapter of the Association for Computational Linguistics (EACL), 2021.
- [20] **Few-shot Intent Classification and Slot Filling with Retrieved Examples**
Dian Yu, Luheng He, Yuan Zhang, **Xinya Du**, Panupong Pasupat and Qi Li
In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2021.
- [21] **QA-Driven Zero-shot Slot Filling with Weak Supervision Pretraining**
Xinya Du, Luheng He, Qi Li, Dian Yu, Panupong Pasupat and Yuan Zhang
In Annual Meeting of the Association for Computational Linguistics (ACL), 2021.
- [22] **Event Extraction by Answering (Almost) Natural Questions**
Xinya Du and Claire Cardie
In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020.
[Top 1% most cited articles published in Computer Science in 2020.](#)
- [23] **Improving Event Duration Prediction via Time-aware Pre-training**
Zonglin Yang, **Xinya Du**, Alexander M. Rush and Claire Cardie
In Findings of the Association for Computational Linguistics: (EMNLP), 2020.
- [24] **Document-Level Event Role Filler Extraction using Multi-Granularity Contextualized Encoding**
Xinya Du and Claire Cardie

In Annual Meeting of the Association for Computational Linguistics (ACL), 2020.

- [25] **Leveraging Structured Metadata for Improving Question Answering on the Web**
Xinya Du, Adam Fourney, Robert Sim, Claire Cardie, Paul Bennett and Ahmed Hassan Awadallah
In Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL/IJCNLP), 2020.
- [26] **Be Consistent! Improving Procedural Text Comprehension using Label Consistency**
Xinya Du, Bhavana Dalvi, Niket Tandon, Antoine Bosselut, Wen-tau Yih, Peter Clark, Claire Cardie
In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2019.
- [27] **Harvesting Paragraph-Level Question-Answer Pairs from Wikipedia**
Xinya Du and Claire Cardie
In Annual Meeting of the Association for Computational Linguistics (ACL), 2018.
Top 1% most cited articles published in Computer Science in 2018.
- [28] **Identifying Where to Focus in Reading Comprehension for Neural Question Generation**
Xinya Du and Claire Cardie
In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2017.
- [29] **Learning to Ask: Neural Question Generation for Reading Comprehension**
Xinya Du, Junru Shao and Claire Cardie
In Annual Meeting of the Association for Computational Linguistics (ACL), 2017.
Featured in New Scientist and TechRepublic [Link]
Top 0.1% most cited articles published in Computer Science in 2017
- [30] **Cornell Belief and Sentiment System at TAC 2016**
Vlad Niculae, Kai Sun, Xilun Chen, Yao Cheng, Xinya Du, Esin Durmus, Arzoo Katiyar, Claire Cardie
In Text Analysis Conference (TAC), 2016.

PROFESSIONAL Chairing:
SERVICES

ACL Rolling Review (ARR) 2024

Annual Meeting of the Association for Computational Linguistics (ACL) 2024 Session Chair
Conference on Empirical Methods in Natural Language Processing (EMNLP) 2024 Area char

Conference on Empirical Methods in Natural Language Processing (EMNLP) Demo Track 2024 Area char

Annual Meeting of the Association for Computational Linguistics (ACL) 2023 Area char
International Conference on Computational Linguistics (COLING) 2024, 2025 Senior Area Chair

Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) 2024 Website chair

Journal Reviewer:

IEEE Transactions on Knowledge and Data Engineering (TKDE)

IEEE Transactions on Audio, Speech and Language Processing (TASLP)

IEEE Transactions on Neural Networks and Learning Systems (TNNLS)

ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP)
ACM Transactions on Knowledge Discovery from Data (TKDD)
Computational Linguistics (CL)
Knowledge and Information Systems (KAIS)
AI Communication

Conference Committee Member:

Annual Meeting of the Association for Computational Linguistics (ACL)
Annual Conference on Neural Information Processing Systems (NeurIPS)
International Conference on Learning Representations (ICLR)
Conference on Empirical Methods in Natural Language Processing (EMNLP)
Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)
Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL/IJCNLP)
International Joint Conference on Artificial Intelligence (IJCAI)
AAAI Conference on Artificial Intelligence (AAAI)
Natural Language Processing and Chinese Computing (NLPCC)
Conference on Computational Natural Language Learning (CoNLL)
Workshop on Noisy User-generated Text (W-NUT)
Workshop on Machine Reading for Question Answering (MRQA)
Joint Conference on Lexical and Computational Semantics

PhD/MS Committee Member:

PhD: Xiangci Li, Jishnu Jaykumar Padalunkal, Basel Abdeen, Yibo Hu.
MS: Wooseong Yang, Shubham Patel

Other Activities:

PhD Admission Committee, UT Dallas, 2022, 2023, 2024
Site host of North American Computational Linguistics Olympiad (NACLO), 2023.
Member of Cornell CS Department PhD Admission Committee, 2021.
Volunteer for Cornell CS Department PhD Visit Day, 2019, 2020, 2021.
Student Volunteer for ACL 2017, ACL 2018, EMNLP 2017.